

Halina Zadora *

EFEKTYWNOŚĆ PREDYKCJI ZMIENNYCH ZAKŁÓCAJĄCYCH

W literaturze rozpowszechniony jest pogląd, że wartości zmiennych objaśniających w okresie prognozowanym można ustalić, budując ich prognozy na podstawie bądź modeli tendencji rozwojowej, bądź też modeli autoregresyjnych. Szczególnie częste, a nawet jedyne staje się to wówczas, gdy zmiennymi prognozowanymi są daty. W pracy zamierzamy wykazać, że prognozowanie dat na podstawie modeli przyczynowo-skutkowych lub symptomatycznych zwiększa błędy predykcji, jeśli wartości zmiennych objaśniających w okresie prognozowanym ustala się na podstawie modeli tendencji rozwojowej, bądź modeli autoregresyjnych. O wiele bezpieczniej i poprawniej jest przyjąć, że zmienna prognozowana jest generowana przez model autoregresyjny. Wówczas prognoza jest funkcją wartości zmiennej prognozowanej, opóźnionej w czasie, które to wartości w okresie prognozowanym są znane.

W dalszej części pracy prognozowanie na podstawie modeli przyczynowo-skutkowych nazywać będziemy I metodą predykcji, a prognozowanie na podstawie modeli autoregresyjnych - II metodą predykcji.

Zapiszmy najprostszy model przyczynowo-skutkowy postaci:

$$(1) \quad Y_t = \alpha X_t + \alpha_0 + \xi_t,$$

gdzie zarówno X_t jak i Y_t są datami. Skoro tak, to zmienna X_t ,

* Dr, adiunkt w Zakładzie Międzynarodowych Stosunków Ekonomicznych w Akademii Ekonomicznej w Katowicach.

jeśli zamierzamy zbudować prognozę Y_T musi być wyjaśniona odrębnym modelem. Przyjmijmy, że

$$(2) \quad X_t = \beta X_{t-1} + \beta_0 + V_t.$$

Podstawiając (2) do (1) otrzymujemy:

$$(3) \quad \begin{aligned} Y_t &= \alpha(\beta X_{t-1} + \beta_0 + V_t) + \alpha_0 + \xi_t = \\ &= \alpha\beta X_{t-1} + (\alpha\beta_0 + \alpha_0) + (\alpha V_t + \xi_t). \end{aligned}$$

Następnie niech

$$\gamma = \alpha \cdot \beta,$$

$$\gamma_0 = (\alpha\beta_0 + \alpha_0),$$

$$Z_t = (\alpha V_t + \xi_t),$$

ponadto niech

$$E(V_t \xi_t) = 0,$$

$$E(X_{t-1} V_t) = 0,$$

$$E(X_t \xi_t) = 0$$

oraz

$$E(\xi_t \xi_{t-1}) = 0 \quad \text{ i } \quad E(V_t V_{t-1}) = 0.$$

Relacje (3) możemy zapisać w następującej prostej postaci:

$$(4) \quad Y_t = \gamma X_{t-1} + \gamma_0 + Z_t,$$

gdzie zarówno Y_t , jak i X_{t-1} , co wynika z (1) i (2) są zmiennymi losowymi, podobnie jak Z_t , której parametry rozkładu są następujące:

$$E(Z_t) = E(\alpha V_t + \xi_t) = 0,$$

$$D^2(Z_t) = E(\alpha V_t + \xi_t)^2 = \alpha^2 D^2(V_t) + 2\alpha E(V_t \xi_t) + D^2(\xi_t) = \\ = \alpha^2 D^2(V_t) + D^2(\xi_t),$$

$$E(Z_t X_{t-1}) = E(\alpha V_t + \xi_t) X_{t-1} = 0.$$

Ponieważ zarówno Y_t , jak i X_{t-1} są zmiennymi losowymi, to równanie (4) uważać należy za funkcję regresji. Parametry tej funkcji wyrażają się wzorami¹:

$$\gamma = \alpha \cdot \beta = \rho_{xy} \frac{D(Y_t)}{D(X_t)} \rho_{xx} \frac{D(X_t)}{D(X_{t-1})},$$

$$(5) \quad \gamma_0 = \alpha\beta + \alpha_0 = \rho_{xy} \frac{D(Y_t)}{D(X_t)} \left[E(X_t) - \rho_{xx} \frac{D(X_t)}{D(X_{t-1})} E(X_{t-1}) \right] + \\ + E(Y) - \rho_{xy} \frac{D(Y_t)}{D(X_t)} E(X_t) = E(Y_t) - \gamma E(X_{t-1}),$$

gdzie

$$\rho_{xy} = E(X_t Y_t) \frac{1}{D(X_t) D(Y_t)},$$

$$\rho_{xx} = E(X_t X_{t-1}) \frac{1}{D(X_t) D(X_{t-1})}.$$

Łatwo wykazać, że

$$(6) \quad D^2(\xi_t) = E[Y_t - (\alpha X_t + \alpha_0 + \xi_t)]^2 = D^2(Y_t) (1 - \rho_{xy}^2),$$

$$(7) \quad D^2(V_t) = E[X_t - (\beta X_{t-1} + \beta_0 + V_t)]^2 = D^2(X_t) (1 - \rho_{xx}^2).$$

Skąd otrzymujemy

¹ Gdzie $\alpha = \rho_{xy} \frac{D(Y_t)}{D(X_t)}$, $\alpha_0 = E(Y_t) - \alpha E(X_t)$ oraz $\beta = \rho_{xx} \frac{D(X_t)}{D(X_{t-1})}$, $\beta_0 = E(X_t) - \beta E(X_{t-1})$.

$$\begin{aligned}
 (8) \quad D^2(Z_t) &= \alpha^2 D^2(V_t) + D^2(\xi_t) = \\
 &= \rho_{xy}^2 \frac{D^2(Y_t)}{D^2(X_t)} D^2(X_t) (1 - \rho_{xx}^2) + D^2(Y_t) (1 - \rho_{xy}^2) = \\
 &= D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) + D^2(Y_t) (1 - \rho_{xy}^2).
 \end{aligned}$$

Oznaczmy wariancję predykcji przez V_1^2 . Otrzymujemy

$$\begin{aligned}
 (9) \quad V_1^2 &= E[Y_T - y_{Tp}]^2 = E[\gamma X_{T-1} + \gamma_0 + Z_T - \gamma X_{t-1} - \gamma_0]^2 = \\
 &= D^2(Z_T).
 \end{aligned}$$

Spróbujmy teraz przyjąć inną procedurę prognozowania. Załóżmy mianowicie, że zmienna losowa Y_t jest generowana przez model autoregresyjny rzędu pierwszego, czyli

$$(10) \quad Y_t = \lambda Y_{t-1} + \lambda_0 + \chi_t.$$

Jeżeli znamy parametry rozkładu zmiennej opisanej relacją (10), to współczynniki λ , λ_0 wyrażają się wzorem:

$$(11) \quad \lambda = \rho_{yy} \frac{D(Y_t)}{D(Y_{t-1})}, \quad \lambda_0 = E(Y_t) - \lambda E(Y_{t-1})$$

oraz

$$(12) \quad D^2(\chi_t) = E[Y_t - (\lambda Y_{t-1} + \lambda_0)]^2 = D^2(Y)(1 - \rho_{yy}^2),$$

gdzie

$$\rho_{yy} = E(Y_t Y_{t-1}) \frac{1}{D(Y_t) D(Y_{t-1})}.$$

Obliczmy wariancję predykcji

$$(13) \quad V_2^2 = E[Y_T - y_{Tp}]^2 = E[\lambda Y_{T-1} + \lambda_0 + \chi_t - \lambda Y_{T-1} - \lambda_0]^2 = D^2(\chi_t).$$

Jak wynika z (9) i (13) wariancja predykcji dla obu metod prognozowania równa jest wariancji resztkowej. Ponieważ uważamy, że jedynie poprawną metodą prognozowania danych jest prognozowanie na podstawie modeli autoregresyjnych wystarczy wykazać, że

$$D^2(Z_t) > D^2(X_t).$$

Ponieważ $D^2(Z_t) = V_1^2$ oraz $D^2(X_t) = V_2^2$, to

$$(14) \quad B = D^2(Z_t) - D^2(X_t)$$

nazywać będziemy obciążeniem I metody predykcji.

Sprawdźmy, jak będą zachowywać się obie wariancje $D^2(Z_t)$ i $D^2(X_t)$, jeśli współczynniki korelacji przyjmą wartości krańcowe. I tak:

1. Jeśli

$$\rho_{xy}^2 = 1, \quad \rho_{yy}^2 = 1, \quad \text{to}$$

$$(15) \quad D^2(Z_t) = D^2(Y_t) (1 - \rho_{xx}^2) = D^2(V_t),$$

$$D^2(X_t) = 0.$$

Wynika z tego, że $B = D^2(V_t)$. Oznacza to, że istnienie funkcyjnej zależności między XY nie chroni nas od błędu predykcji, którego wartość równa jest wariancji resztkowej $D^2(V_t)$. Obie procedury prognozowania są równoważne dopiero wówczas, gdy $\rho_{xx}^2 = 1$, tzn. gdy występuje zależność funkcyjna również między zmiennymi X_t i X_{t-1} .

2. Jeśli

$$\rho_{xy}^2 = 0 \quad \text{i} \quad \rho_{yy}^2 = 0, \quad \text{to}$$

$$(16) \quad D^2(Z_t) = D^2(Y_t),$$

$$D^2(X_t) = D^2(Y_t).$$

Oznacza to, że obie procedury prognozowania są równoważne w sensie efektywności, a błąd predykcji nie zależy od siły korelacji między zmiennymi X_t X_{t-1} , bowiem $B = 0$ dla dowolnych ρ_{xx}^2 .

O wiele ciekawsza jest analiza błędów predykcji obu metod prognozowania, gdy przyjmiemy, że $\rho_{xy}^2 \neq 0$ $\rho_{yy}^2 \neq 0$ $\rho_{xx}^2 \neq 0$. Przy takich wartościach współczynników korelacji mogą zachodzić następujące sytuacje:

Sytuacja 1, gdy $\rho_{yy}^2 = \rho_{xy}^2$

Wówczas

$$B = D^2(Z_t) - D^2(X_t) = D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) + D^2(Y_t) (1 - \rho_{xy}^2) - \\ - D^2(Y_t) (1 - \rho_{yy}^2) = D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2).$$

Wykorzystując relację (7) otrzymujemy ostatecznie, że

$$(17) \quad B = \frac{D^2(Y_t)}{D^2(X_t)} \rho_{xy}^2 D^2(V_t) = \alpha^2 D^2(V_t).$$

Skąd wynika, że

$$B > 0.$$

W n i o s e k: Jeżeli zmienne objaśniające w obu modelach są w jednakowym stopniu skorelowane ze zmienną prognozowania Y_t , to I metoda predykcji jest gorsza pod względem efektywności, bowiem wariancja predykcji V_1^2 jest większa o $\rho_{xy}^2 D^2(V_t)$ od wariancji V_2^2 . Obie procedury prognozowania są równoważne tylko wówczas, gdy $\rho_{xx}^2 = 1$.

Sytuacja 2, gdy $\rho_{yy}^2 - \rho_{xy}^2 > 0$, czyli $\rho_{yy}^2 > \rho_{xy}^2 \neq 0$

Wówczas

$$B = D^2(Z_t) - D^2(X_t) =$$

$$\begin{aligned}
&= D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) + D^2(Y_t) (1 - \rho_{xy}^2) - D^2(Y_t) (1 - \rho_{yy}^2) = \\
&= \alpha^2 D^2(V_t) + D^2(Y_t) (1 - \rho_{xy}^2 - 1 + \rho_{yy}^2) = \\
&= \alpha^2 D^2(V_t) + D^2(Y_t) (\rho_{yy}^2 - \rho_{xy}^2).
\end{aligned}$$

Zauważmy, że

$$K = \rho_{yy}^2 - \rho_{xy}^2 > 0.$$

Wobec tego

$$(18) \quad B = \alpha^2 D^2(V_t) + K D^2(Y_t) > 0.$$

W n i o s e k: Jeżeli stopień autokorelacji zmiennej Y_t jest wyższy od poziomu korelacji tej zmiennej ze zmienną X_t , to efektywność predykcji pogarsza się w porównaniu do sytuacji 1. Zauważmy, że nawet przy maksymalnej wartości ρ_{xx}^2 I metoda jest mniej efektywna niż metoda II. Jeżeli więc $\rho_{yy}^2 > \rho_{xy}^2$, to zawsze I metoda predykcji daje gorsze prognozy, bowiem obarczone są większym błędem.

Sytuacja 3, gdy $\rho_{yy}^2 - \rho_{xy}^2 < 0$, czyli $\rho_{yy}^2 < \rho_{xy}^2$

Wówczas

$$\begin{aligned}
(19) \quad B &= D^2(Z_t) - D^2(X_t) = \\
&= D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) + D^2(Y_t) (1 - \rho_{xy}^2) - D^2(Y_t) (1 - \rho_{yy}^2) = \\
&= D^2(Y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) + D^2(Y_t) (\rho_{yy}^2 - \rho_{xy}^2).
\end{aligned}$$

Sprawdźmy czy zawsze $B > 0$. Zauważmy, że $B > 0$ wtedy, gdy

$$(20) \quad D^2(y_t) \rho_{xy}^2 (1 - \rho_{xx}^2) > D^2(y_t)(\rho_{yy}^2 - \rho_{xy}^2).$$

Z (20) wynika, że

$$\rho_{xy}^2 (1 - \rho_{xx}^2) > \rho_{yy}^2 - \rho_{xy}^2.$$

Podzielmy obie strony nierówności przez ρ_{xy}^2 . Otrzymujemy, że

$$1 - \rho_{xx}^2 > \frac{\rho_{yy}^2}{\rho_{xy}^2} - 1,$$

a następnie

$$(21) \quad 2 - \rho_{xx}^2 > \rho_{yy}^2 : \rho_{xy}^2.$$

Ponieważ z założenia $\frac{\rho_{yy}^2}{\rho_{xy}^2} < 1$, to nierówność (21) będzie zawsze

spełniona. Wykazaliśmy tym sposobem, że jeżeli $\rho_{yy}^2 < \rho_{xy}^2$, to zawsze $B > 0$.

Wniosek: W sytuacji, gdy intuicyjnie wydawałoby się, że prognozowanie na podstawie modelu (1) powinno dawać nie gorsze prognozy niż na podstawie modelu (2), okazuje się jeżeli $\rho_{yy}^2 < \rho_{xy}^2$, to zawsze I metoda predykcji jest obarczona większym błędem.

Nasze rozważania prowadzone były przy bardzo ostrym założeniu, że nie występują błędy szacunku nieznanymi parametrów rozkładu dwumiarowej zmiennej losowej. Założenie to wprowadzono celowo, bowiem bardzo utrudniłoby to dowód, a nie zmieniłoby ogólnego wniosku o mniejszej efektywności I metody prognozowania. Ponadto dowód ten nie jest ogólny, jako że rozważaliśmy najprostsze postacie funkcji regresji. Uogólnienie na n zmiennych objaśniających nie jest niemożliwe i można już na podstawie powyższych wyników domniemywać, że efektywność atakowanej metody predykcji będzie maleć na skutek występowania współczynników autokorelacji $\rho_{x_{it} x_{it-j}}$, gdzie $i = 1, 2, \dots, n$ jest ilością zmiennych, a $j = 1, 2, \dots, p$ jest opóźnieniem tych zmiennych.

Halina Zadora

THE EFFICIENCY OF PREDICTION OF NOISE VARIABLES

The author investigates the efficiency of prediction of noise variables in two cases, namely when these variables are generated by causal models and by autoregressive models. In the first case as the values of (random) explanatory variables in the horizon of prediction are not known, it is assumed that their extrapolated values are used instead. It is proved that the efficiency of such prediction is always smaller than the efficiency when autoregressive models are used.